Minutes for the P-20 Subcommittee on Teacher Effectiveness
May 5, 2010

Attendees

Josh Anderson
Xian Barrett
Bette Bergeron
Bill Chafetz
Ann Courter
Vicky Chou
Lizanne DeStefano
Erika Hunt
Jim O'Connor
David Osta, New Teacher Center
Michelle Seelbach
Lisa Hood, Leadership Effectiveness Committee, ISU
Audrey Soglin
Brad White

Minutes:

Josh reviewed the four areas that this subcommittee will focus on: certification and recertification; a data dashboard for human capital; a survey to assess the relationship between teacher and supervisor; a predictive selection model for teachers (a 4-5 year project).

 Josh reviewed Article 21 and referred to a document titled "The Current State of Teacher Education Program Entry and Certification Requirements." Rule making will be the place where we can make an impact with our recommendations for teacher certification and re-certification.

Discussion of tests for entry and certification: Basic Skills test
Audrey stated that she was at a presentation by The New Teacher Project (TNTP) at the Performance Advisory Council about tests of academic aptitude as a predictor of student growth. TNTP had stated that there wasn't a connection between aptitude and later teacher effectiveness.

Josh and Jim offered to put together a summary of the research on how teacher academic aptitude relates to later student achievement. The range of the teacher scores studied is important--there is a difference when you compare teachers with ACT scores of 10 and 30, but less difference between teachers with ACTs of 20 and 23.

David Osta noted that the research is mixed in this area, and stressed that more than one factor is needed to predict success as a teacher—academic aptitude is only one indicator.

Vicki Chou referred to the McKinsey study, and noted that to have a teaching force from the top third of college graduates, like Singapore and Finland, we would need 40% of them to enter teaching. Bill noted that we'd have to have significant change in society to get people to choose teaching at a low performing school over working in an investment bank. A question was raised whether the McKinsey study actually studied the impact of academic aptitude on teaching, or just reported practices in countries.

There was a discussion about the timing of basic skills testing—the results could be very different if the test is given to freshmen or sophomores rather than at college graduation. The later it is given, the more time students have to gain mastery of the areas tested in the basic skills test, and the more time the colleges have to help them progress.

Brad asked for everyone's sense of where the cut score should be relative to the ACT. Vicki said that her researchers saw it at 21 for reading. An ACT score of 26 represents the top third.

Brad noted that a North Carolina study saw that the extremes were predictive: below 18, and above 27. He noted that we don't know whether the candidates who fell below a certain threshold would have been good teachers because their scores kept them out of the profession. In response to Audrey's question about whether content area matters, Brad noted that academic aptitude is more predictive for math instruction than for successful reading instruction.

Bill explained that the Chicago Public Education Fund has looked at basic skills test scores from 240 and above, and also down to 100, and compared scores to student valued-added scores in reading and math, but only for one year, and the data was not stable.

It was noted that high performers may not score particularly high on the basic skills test if it is seen as easy to cross the threshold, and a higher score does not matter.

David suggested that we want the basic skills test to be the first filter, and not so fine-grained that it will create false negatives, blocking good people. There are later filters: the content tests, hiring, tenure, evaluations, etc.

Josh noted that in CPS we have 25,000 applicants for 2,000 spots. What do we have to lose by raising the cut score? David responded that in places where you have an over-supply, you should push the number higher, but not all parts of the state have an over-supply.

Audrey disagreed with the proposition that in order to make teaching attractive to high performers the bar must be set high, noting that there are other more important factors that make the profession attractive or less attractive.

Bill asked whether anyone has done a value-added supply chain analysis? If university training can add significant skills and aptitudes to the candidates, then you don't need a good filter at entry to university programs.   Any test of cognitive ability frequently comes with adverse impact, as shown by decades of industrial psychology research. If teaching is a learned profession and we create strong apprenticeships, perhaps academic aptitude is not the best first screen. Fit/disposition could be a better first gate, and more suggestive of future trajectory.

Lizanne noted that teacher ed programs have been concerned that it would be unfair to students to allow them to spend tuition and time in a teacher ed program only to fail the basic skills test later. Bette noted that NCATE accreditation has required earlier testing since 2002. Audrey pointed out that there are many other screens in addition to the basic skills test along the way to completing teacher certification, including course grades and field experience. Students may fail to make it through those screens after spending considerable time and tuition, too.

Assessment of Professional Teaching (APT test)
The test currently is tailored for different grade levels and subject areas, and generally taken after student teaching. The pass rates are about 94%. Teach for America teachers take the exam after the first year of teaching.   No one was aware of any research done on the APT or the content tests.

Other assessments of teaching, particularly for math and science, are being developed. Illinois is piloting the Teacher Performance Assessment (TPAC) from California at UIC and 3 other schools. The test includes video-taping students teaching 3-5 lessons, and evaluated by trained scorers. Vicki explained that a UIC professor liked the test so much she created a formative mid-term version of the test, and students have found the process of completing the TPAC to be valuable.

Josh compared the test to the simulation that medical students are required to undergo, diagnosing ailments in actors. David commented that having someone do actual teaching is really helpful to determine whether someone will be successful.

Content Tests
The research on content tests shows a stronger relationship between content knowledge in science and math and later students' achievement than in other subjects. There is a better tie between later student performance and content tests than student performance and aptitude tests. Pass rates are around 94%. The programs prepare their students to pass it—something they cannot do with basic skills tests given to enter the program.

The new Common Core standards may require updating the content tests.

Disposition

Bette stated that programs that are NCATE accredited need to assess disposition. At various times a student takes a disposition self-assessment. It can also be a red flag in field experience. Vicki noted that UIC uses the online Haberman test. UIC faculty have also created an interview instrument that links dispositions to behavior. Disposition tests measure such qualities as persistence, grit, emotional intelligence, perseverance, personal responsibility, leadership, interpersonal skills, and understanding the perspectives of others.

CPS is doing an assessment with Polaris for disposition. Gallup and others also have tests. They screen out the bad, rather than assess degree of future success—a left-tail test. There was discussion about whether dispositions are stable over a lifetime, when a test should be taken, how stable personality. The question was raised whether these are characteristics that should be assessed at hiring, rather than by a state test.

Wrap-up
 At each testing point, we currently think about passing the test as a threshold. As an organization, TFA looks to use a weighted combined score. Could we look at a blended or holistic model? Xian noted that there would be merit in treating the testing as diagnostic, giving teachers the right mindset that they will be growing and developing as professionals.

It was suggested that perhaps we could look at different basic skills thresholds for different grades and subjects taught.

Lizanne noted that we need to be careful about what a state level assessment can do. Some of these things need to be built in the teacher ed programs themselves.

The meeting was adjourned with additional agenda items to be discussed at future meetings.